Springer Nature 科研数据政策标准化工作实践及启示*

■ 王丹丹　刘清华　葛力云
河南科技大学管理学院　洛阳 471023

摘　要：[目的/意义]Springer Nature 在推动科研数据政策标准化和统一化发展方面做出了开拓性贡献，对 Springer Nature 提出的科研数据政策标准框架及其推广应用情况进行研究，以期为我国科技期刊管理机构出台适合我国科技期刊的科研数据政策标准框架并推广实施提供参考和借鉴。[方法/过程]通过网络检索、文本分析、网站访查和对比分析等方法，对 Springer Nature 的政策标准框架进行解析，对 Springer Nature 为推行政策标准框架而提供的支持服务进行研究，系统梳理和总结 Springer Nature 推进科研数据政策标准化的工作实践，揭示其借鉴意义。[结果/结论]Springer Nature 的政策标准化框架可以作为我国相关部门制定科研数据政策标准框架的参考，Springer Nature 为推行政策标准框架而开发的第三方监管服务，展示了出版商这一群体参与科研数据管理服务的实践路径。建议由我国科技期刊管理机构牵头，召集利益相关群体，借鉴国际先进经验建构科学合理的政策框架，并设计配套支持服务以有效推动政策框架的应用。

关键词：出版商　科研数据　政策　服务
分类号：G231
DOI：10.13266/j.issn.0252-3116.2020.18.015

1　引言

数据密集型科学环境下，科技创新越来越依赖于科研数据的获取和分析，科研数据存储和共享的重要性日益凸显。科研数据政策会影响科研人员在不同程度上共享数据的意愿[1]，越来越多的资助机构发布有关科研数据共享的政策，其中包括美国国立卫生研究院（National Institutes of Health，NIH）、盖茨基金会（Gates Foundation）、欧盟2020年地平线计划（EU Horizon 2020 programme）、惠康基金会（Wellcome Trust）以及七个英国研究委员会（UK research councils）。这些机构的政策要求各不相同，有的要求科研人员准备科研数据管理计划，而有些则要求将数据存档的证据在已发表的论文中提供，例如工程和物理科学研究委员会（Engineering and Physical Sciences Research Council，EPSRC）。

2015年以来，越来越多的期刊和出版商认可并开始实施“透明度和开放度提升（Transparency and Openness Promotion，TOP）”指南。TOP指南是一套全面但理想的期刊政策，包含八个模块化标准[2]。期刊和出版商把完全遵从TOP指南作为长期奋斗目标，通过调整政策，不断接近这一目标。已有研究表明，出版商和期刊通过制定并不断完善其政策，可以帮助提高科研人员对数据共享问题的认识[1]。虽然出台科研数据政策的出版商和期刊的数量越来越多，但是由于他们多是从各自研究社区的需求和规范出发制定，致使这些政策在内容、可发现性、易于解释性、基础架构集成以及对合规性的支持等方面各不相同，甚至存在较大差异[3]。由于政策缺乏统一的标准，对于期刊编辑而言，制定和支持数据政策非常困难；对科研人员而言，理解和遵守数据政策也不容易；对基础设施提供商和科研支持人员而言，协助期刊作者遵守政策也变得异常复杂[4]。

这使得标准和相对统一的科研数据政策研究成为

* 本文系河南省高等学校哲学社会科学研究优秀学者资助项目（项目编号：2019-YXXZ-09）和河南科技大学大学生研究训练计划“基于期刊的社会科学科研数据特征分析及其启示”（项目编号：2019250）研究成果之一。
作者简介：王丹丹（ORCID：0000-0001-9214-4906），教授，博士，E-mail：wangdan430806@163.com；刘清华（ORCID：0000-0003-2350-8061），硕士研究生；葛力云（ORCID：0000-0002-3774-4615），硕士研究生。
收稿日期：2020-01-17　修回日期：2020-03-30　本文起止页码：137-145　本文责任编辑：杜杏叶

新热点。在化学、高能物理和医学领域,出现了针对特定学科的举措,以标准化和统一期刊和出版商的科研数据政策。2017年国际医学期刊编辑委员会(International Committee of Medical Journal Editors,ICMJE)为其成员期刊(包括 *BMJ*,柳叶刀(*Lancet*),*JAMA* 和《新英格兰医学杂志》)引入了标准化的数据共享政策[5]。ICMJE 的政策没有强制要求数据共享,但却要求必须提供数据可用性声明。数据可用性声明是期刊和出版商数据政策中的常有元素。*PLOS ONE*,*Nature* 和 *BMC* 都要求提供数据可用性声明[6]。一些科研资助机构,如英国七个研究委员会,也要求在已发表的文章中提供数据可用性声明[7]。数据可用性声明也是英国研究委员会(Research Councils UK,RCUK)数据政策通用原则10的要求。已有研究表明反映在投稿指南中的期刊政策有助于提高作者研究的透明度[8-9],影响因子越高的期刊往往采用的数据共享政策越强硬[10],强硬的数据政策不仅要求核验作者数据共享的情况,而且要求提供数据可用性声明。

Springer Nature 在政策标准化方面进行了有益的实践探索,引入了科研数据政策标准框架,一方面为期刊根据自身情况快速制订遵循最佳实践且能够有效推行的数据政策提供了参考和依据,另一方面也为利益相关群体提供了交流的平台[11]。与此同时,为了配合框架的实施,还推出了数据可用性指南并开发了第三方数据监管服务。Springer Nature 推进政策标准化工作的成效显著。笔者同步调研了 Elsevier,Wiley,PLOS,Taylor & Francis 等一些国外出版商,以了解其在期刊政策标准化方面的实践进展,发现这些出版商也都已经意识到了标准化工作的重要性,也有类似的举措,全球科研数据联盟兴趣小组还开发了一个政策框架,用于协调所有出版商之间的数据政策[12]。但与 Springer Nature 相比,这些出版商的标准化工作实践在系统性、完整性方面还存在一定差距,框架推广应用的成效也没有 Springer Nature 好。鉴于此,通过对 Springer Nature 的政策标准框架进行解析,对 Springer Nature 为推行框架而提供的支持服务进行揭示,系统梳理和总结 Springer Nature 推进科研数据政策标准化的工作实践,以期为我国科技期刊管理机构出台适合我国科技期刊的科研数据政策标准框架并推广实施提供参考和借鉴。

2 Springer Nature 科研数据政策标准框架

2.1 提出背景

Springer Nature 注意到其旗下仍有许多期刊没有关于科研数据的明确政策,有政策的期刊也存在政策的强制性和选择性模棱两可,如何执行政策缺乏清楚说明等问题。它认为期刊和出版商有责任支持科研人员遵守资助机构的政策(无论是开放获取还是开放数据)。除遵守政策外,使链接到学术出版物的数据共享,有助于增加论文的可见度,提高论文的引用次数,有助于出版商出版更为丰富的内容,改善对数据的访问和理解,从而推动进一步的研究,晶体学、遗传学、考古学和语言学领域的实践已充分证明了这一点[13]。为了获得这些益处,解决复杂多样的期刊政策问题,Springer Nature 于2016年推出了适用于其出版集团旗下2 500多种期刊的科研数据政策标准框架[14]。其旗下的每种期刊在制定和优化科研数据政策时均可依据此框架,并根据各自所属学科和社区的特点进行个性化调整。

Springer Nature 提出该框架的目的是:①通过该框架让期刊在制定和优化政策时有据可依,既保障统一性,又兼顾个性化;②帮助科研人员达到科研机构和资助机构数据共享的要求;③确保所有的期刊和作者均能按照学科领域最佳的数据共享和归档实践发表成果。为了实现上述目标,Springer Nature 首先召集组建了专家工作组,咨询访谈了涵盖所有学科的30余位期刊编辑和其他一些资助机构、图书馆员后,讨论确定 Springer Nature 期刊科研数据政策的9个共有要素,即:支持通过存储库共享数据;允许引用数据;出版商帮助台;针对特定类型的数据,进行公共数据存放和数据集标识符检查;数据可用性声明;要求并检验公共数据存放和数据集标识符;数据引用规范;数据同行评审;集成数据存储库。其次,针对9个要素咨询并广泛听取期刊编辑、图书馆员和科研资助机构的意见。最后,考虑9个要素由于期刊和学科而导致的适用情形的差异性,提出了四种类型的政策模板,见表1。无论隶属哪个学科,期刊均可结合其学科社区可用资源状况,选取最适合其需求的政策类型。

表1 通用政策框架的四种类型政策模板[11]

四种类型	适用情形
类型一	鼓励数据共享和数据引用
类型二	鼓励数据共享和研究材料共享
类型三	鼓励数据共享,并要求提供数据可用性声明
类型四	要求共享数据及研究材料;共享的数据需进行同行评审;并提供可用性声明

2.2 框架内容

各个期刊和研究团体(research communities)可能

王丹丹，刘清华，葛力云. Springer Nature 科研数据政策标准化工作实践及启示[J]. 图书情报工作,2020,64(18):137－145.

会处于准备支持数据共享的不同阶段,有的已强制要求开放数据,有的则刚开始讨论是否要共享数据这一问题。之所以给出四种类型政策模板,就是在尊重发展阶段差异性的同时提供一个通用且易于理解的框架,以鼓励跨所有研究领域的良好数据共享实践。四种类型政策的对比见表 2[11]。9 个要素中,四种政策都对“支持通过存储库共享数据”和“允许引用数据”两个要素做出了强制性规定。类型一、类型二要求相对宽松,给予科研人员较大的自主权,类型三、类型四更为强硬,主要面向价值较高的数据,通过强制要求确保数据最大程度的开发、重用。从类型一到类型四数据政策强硬程度逐渐递增。类型一鼓励数据共享和数据引用,并为科研人员提供数据存储库列表;类型二建议提供有关准备数据可用性声明的信息;类型三,要求提供数据可用性声明,类型四要求开放数据,并要求数据进行同行评审。

表 2　四种类型政策在 9 个共有要素上的体现

要素	解释	类型一	类型二	类型三	类型四
1. 支持通过存储库共享数据 Data sharing via repositories supported	通过存储库共享的详细信息在作者的期刊指南中有所提及	强制	强制	强制	强制
2. 允许引用数据 Data citation permitted	指南允许作者引用参考列表中的公开可用数据集	强制	强制	强制	强制
3. 出版商帮助台 Publisher helpdesk	作者期刊信息中包含的服务台联系方式	建议	建议	建议	建议
4. 针对特定类型的数据,进行公共数据存放和数据集标识符检查 Public data deposition and dataset identifier checks for specific types of data	在那些已经建立起强制性要求的社区,将检查数据存储作为出版过程的一部分	无要求	建议	强制	强制
5. 数据可用性声明 Data availability statements	在发表文章中的声明,说明如何支持数据可以访问	无要求	建议	强制	强制
6. 要求并检验公共数据存放和数据集标识符 Public data deposition and dataset identifier required and verified	公开提供的数据和所有已发布文章的数据标识符(敏感/个人数据除外)	无要求	无要求	建议	强制
7. 数据引用规范 Data citations	提供并验证参考列表中的相关数据集引用	无要求	无要求	建议	强制
8. 数据同行评审 Peer review of data	同行审稿人指南和流程提供有关访问和查看数据文件的指导	无要求	无要求	建议	强制
9. 集成数据存储库 Integrated data repository	提交系统/审查过程与具体的期刊或通用存储库集成,比如 figshare	无要求	无要求	建议	强制

2.3　要素解析

为了支持框架应用,Springer Nature 于 2016 年 7 月推出了科研数据服务台(helpdesk)[15]。服务台向期刊编辑、作者(科研人员)和支持人员提供基于电子邮件的免费建议,以帮助他们实施和遵守期刊数据政策。M. Astell 等的研究发现,截至 2017 年 10 月,服务台已收到 300 多个查询。通过服务台寻求帮助的人中,超过一半(53%)是科研人员(作者);43% 是期刊编辑,其余 4% 是其他利益相关者,如图书馆员和数据存储库管理者。查询最多的四类问题依次是实施政策(122 个)、数据存储库(53 个)、遵守政策(46 个)、撰写数据可用性声明(31 个)[16]。鉴于此,Springer Nature 在其网站上专门整理提供了数据存储库列表[17]和数据可用性声明指南[18]。本部分以数据可用性声明这一要素为例,对 Springer Nature 的具体实践进行揭示。

数据可用性声明是由作者编写,旨在提供有关在何处可以找到支持其文章中报告的结果的数据以及是否和如何获得这些数据的信息。尽管这些声明没有强制性格式,但是期刊一般应给作者提供模板示例。2016 年 9 月,Nature 宣布所有已接受在《自然》杂志上,以及自然的生命科学期刊上发表的论文,必须提供数据可用性声明[19]。通过引入数据可用性声明,使《自然》的政策与 Springer Nature 的政策标准框架保持一致。

Springer Nature 政策框架中,类型二、类型三和类型四均鼓励或强制要求提供数据可用性声明。不同类型数据的开放程度不一样,有些数据可以开放获取,而涉及到隐私或机密的数据则需要经过许可才能获取,因此,不同类型的数据应该有自己独特的“可用性声明”。一些常见的可用性声明如:①在当前研究期间和/或分析期间生成的数据集可在 *** 存储库中获得(The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]);②在当前

研究期间产生/或分析的数据集可以在合理的请求下从相应的作者处获得(The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request);③本研究期间生成或分析的所有数据均包含在本文(及其补充信息文件)中(All data generated or analysed during this study are included in this published article (and its supplementary information files));④在当前研究期间和/或分析期间生成的数据集由于[数据不公开的原因]而不公开的情况下,可通过合理请求从相应作者处获得(The datasets generated during and/or analysed during the current study are not publicly available due to [REASON(S) WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request);⑤数据共享不适用于本文,因为在当前研究期间没有生成或分析数据集(Data sharing not applicable to this article as no datasets were generated or analysed during the current study);⑥支持本研究结果的数据可从[第三方名称]获得,但这些数据的可用性受到限制,它们只可在当前研究的许可下使用,因此不公开。但是,根据合理要求并经[第三方名称]许可,作者可获得数据(The data that support the findings of this study are available from [THIRD PARTY NAME] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [THIRD PARTY NAME].)。如果期刊没有关于数据可用性声明的明确要求,作者可以使用或调整上述表述,以确定最适合其论文的可用性声明。根据研究的性质,有时可能需要结合几种说法。Springer Nature 提供了期刊文章中支持数据可用性声明的示例,如表 3 所示:

表 3 期刊文章中支持数据可用性声明示例

声明类型	模板/示例文本
研究期间生成的数据受数据共享授权的约束,并且可以在不使用 DOIs 发布数据集的公共存储库中使用	支持该研究结果的[数据类型,如"序列"]数据已存放在[存储库名称,如"genbank"]中,并带有[主要]登录码[列出链接的登录码,如"kp253039"]
使用 DOIs(非强制存储)发布数据集的公共(机构、一般或特定主题)存储库中可用的数据	支持该研究结果的[数据类型]数据可在[存储库名称,例如"figshare"]中找到,标识符为[数据 DOI],例如"http://doi.org/10.6084/m9.figshare.1499292"[参考号]
在没有使用 DOIs(非强制存储)发布数据集的公共(机构、一般或特定主题)存储库中可用的数据	支持该研究结果的[数据类型,如" Snow pack depth "]数据可从[存储库/资源名称,如" GlobSnow "]、[超链接到数据集/数据源,例如"http://wwwGuangs.If/"]中获得
作者可以确认所有相关数据都包含在文章和/或其补充信息文件中	作者声明,支持本研究结果的[所有其他]数据可在文章[及其补充信息文件]中找到
由于隐私或其他限制,数据可根据作者要求提供	在合理的要求下,支持该研究结果的数据可根据相应作者[作者首字母]的要求提供。由于[国家限制,如"包含可能危及研究参与者隐私/同意的信息"],这些数据无法公开获取
受第三方限制的数据	支持该研究结果的数据可从[第三方名称]获得,但这些数据的可用性受到限制,它们只可在当前研究的许可下使用,因此不公开。然而,在合理的请求下,并在得到[第三方名称]的许可下,数据可从作者处获得
从公共资源派生并随文章提供的数据集	当前研究期间分析的数据集可在[存储库名称]存储库[数据标识符,例如 doi:10.7910/DVN/HEWGDD][参考编号]中找到。这些数据集源自以下公共域资源:[列表资源及其 URL]
中央大型设施中生成的数据:一些物理科学中的常见实践	原始数据是在[设施名称]大型设施中生成的。根据要求,可从相应的作者处获得支持该研究结果的衍生数据
不适用数据共享——例如,如果一篇文章完全描述理论研究,无新数据	①数据共享不适用于该文,因为当前研究期间没有生成或分析数据集。②如果没有使用辅助数据源,也没有创建新数据,但需要提供数据获取声明,则可以提供以下说明:"在研究期间没有创建新数据。"
公开可用数据	①该研究中创建的所有数据都可以从巴斯大学研究数据档案室(在此处插入 DOI)中公开获取。②支持本研究的所有数据均作为本文的补充信息提供。③本文的结果完整地提供了所有数据
引用多个数据集	①您可以在本文的参考资料部分引用多个数据集,可以像引用论文一样引用数据;您还可以创建一个单独的存档记录,该记录提供所有数据集的链接(如果这些数据集存档在不同的位置)。②此出版物受多个数据集支持,这些数据集在参考部分的引用位置公开可用
数据的二次分析	①该研究是对现有数据的重新分析,这些数据可以在参考资料部分的引用位置公开获得。有关数据处理的文件可在巴斯大学科研数据档案库[DOI]找到。②该研究汇集了根据要求获得的现有数据,并受到来自多个不同来源的许可证限制。有关如何获取这些数据集的详细信息,请参阅[DOI]中提供的文档

(续表 3)

声明类型	模板/示例文本
敏感数据	①如果存在不公开共享数据的法律或道德理由,应在数据访问声明中提供此理由,并提供有关数据是否完全受限制或在特定条件下是否可以被访问。②出于道德考虑,无法公开提供支持数据。有关数据和访问条件的更多信息,请访问巴斯大学科研数据档案库[DOI]。③由于研究的(商业、政治、伦理)敏感性,没有受访者同意保留或分享他们的数据。有关该数据其他方面的更多详细信息,可从巴斯大学科研数据档案库获取 [DOI]。④英国数据服务中心[DOI]提供的支持数据可供真正的科研人员使用,但需要注册
商业限制	①如果在共享数据方面有协议、未决专利或其他合同限制,应在数据获取声明中提供这些信息。②由于与研究合作者签订了保密协议,只有在遵守保密协议的情况下,才能向真正的科研人员提供数据。有关数据和如何请求获取的详细信息,请访问巴斯大学科研数据档案库[DOI]
非数字数据	①如果使用二手数据或第三方拥有的数据,则不太可能存档或共享原始数据。使用二手数据时,应提供有关数据来源和访问安排的信息。如果您正在使用多个数据源,应该考虑提供一个文档,总结所有数据源、访问方式以及数据处理方法等。该研究是对现有数据的重新分析,这些数据可以在参考资料部分的引用位置公开获得。有关数据处理的进一步文件可在巴斯大学科研数据档案库[DOI]找到。②支持该研究的非数字数据由巴斯大学的相应作者存储。有关如何获取这些数据的详细信息,请参阅位于[DOI]巴斯大学科研数据档案库中的文档

数据可用性声明是一种简单且可互操作的机制,用于在利益相关者、出版平台和研究学科之间传达支持数据的可用性,它有助于资助机构、出版商和研究社区监督数据政策的遵守情况。可以预见,未来数据可用性声明的普及程度会不断提高。然而,在录用的文章中增加数据可用性声明会增加编辑处理稿件的时间,R. Grant 等的研究发现,对每篇文章添加数据可用性声明,会使专业编辑的稿件处理时间平均增加 10 分钟。但与数据可用性声明的重要性和益处相比,增加的时间应该是可以接受的[20]。对期刊编辑部而言,在提出数据可用性声明要求时,数据共享实践的学科差异一定要考虑,例如在生命科学领域有大量社区存储库,而且生命科学社区的数据共享要求长期存在,因此在《自然》杂志上要求可用性声明能够得到顺利执行。但是在高能物理学领域中,生成和分析的数据可能非常庞大,使得在线共享数据具有挑战性。在这种情况下,“根据合理要求提供数据”可能是一个务实的选择。

3 Springer Nature 配合框架实施的配套支持服务

3.1 服务概况

为配合框架实施,Springer Nature 同步推出了配套支持服务,重点是监管支持同行评审出版物的数据集(Springer Nature 将数据集定义为支持研究项目、研究或出版物的文件的集合,数据集可以包含多个文件,具有多种文件类型),同时兼顾到期刊编辑、同行评审专家、读者以及作者(数据创建者)的各种需求。按服务对象,可分为面向科研人员的服务和面向期刊的服务,见表 4[21]。科研人员和期刊编辑以邮件方式发送服务请求至固定邮箱。收到邮件后,Springer Nature 会在两个工作日内与其联系。按是否收取费用,可分为免费数据服务(面向科研人员和所有期刊编辑)和付费数据服务(只面向科研人员)。与此同时,Springer Nature 一方面鼓励科研人员在 *Scientific Date* 和 *BMC Research Notes* 等数据期刊发表数据;另一方面建立越来越多的科研数据共享学习与交流社区(如数据对话、资源、活动等讨论社区[22]),为科研人员提供互动平台,科研人员可以在社区中了解有关科研数据的最新信息。

表 4 Springer Nature 提供的配套支持服务

Springer	免费数据服务	付费数据服务
服务内容	①面向科研人员,提供政策建议和咨询。咨询内容包括数据政策、数据知识库选择、可用性声明编写、正确引用数据以及其他 ②面向期刊编辑,提供政策框架建议及政策制定咨询服务	①上传数据集到 figshare 托管的安全在线数据储存库 ②专业科研数据编辑人员检查数据文件的完整性,敏感数据,以及数据与相关手稿的一致性 ③通过元数据编辑使数据更容易被发现、获取、互操作和重用 ④将数据出版与作者论文出版进行协调

3.2 第三方数据监管服务

为了更多地了解科研人员在共享数据时面临的实际挑战,Springer Nature 在 2018 年发布了一项针对科研人员的调查,该调查收到 7 700 多个回复。受访者认为,数据共享的主要障碍是“以一种可展示和有用的方式组织数据”(46%)。其他常见的挑战是“不确定版权和许可”(37%),“不知道要使用哪个存储库”(33%),“没有时间来存储数据”(26%)和“共享数据的成本”(19%)。最常报告的障碍表明,希望整理、共享、管理和存档数据的科研人员可能缺乏技能、支持或资源。对资助机构(n = 13)和高等教育机构(n =

15）的访谈结果显示元数据技能和专业知识通常被认为是数据共享的障碍[23]。这一结果与出版商 Wiley、Elsevier 先前的调查结果相类似[24-25]。Springer Nature 认为科研人员应该有更快、更轻松的数据存储途径。但是把科研人员培养成数据管理专家不应是 Springer Nature 的目标，在 Springer Nature 看来，与利益相关群体密切合作，寻找简化数据存储工作流程的方案，开发第三方监管服务（收费服务）才是成功之道。Springer Nature 引入独立于期刊编辑部，由第三方机构提供的付费监管服务的目的是：①方便科研人员更轻松地满足出版商、所属机构和资助机构的数据共享要求；②提升数据的可发现性和可重用性；③减少科研人员用于数据存储、共享和管理的时间；④使数据能够以科学的方式呈现出来；⑤提升科研人员的可见度、影响力，增加研究成果使用的可能性。

2017 年，Springer Nature 推出了一项试点服务[24]，以为希望提供其数据的科研人员提供更多支持。这项服务为科研人员提供手把手的帮助，帮助他们将数据上传到存储库，选择适当的许可，增强元数据，交叉引用数据和任何关联的出版物。数据在 Springer Nature 的 figshare 存储库进行托管和保存。该服务最初可供提交文稿以选择 BMC 期刊的作者使用，随后使用范围扩展到近 100 种期刊，从 2018 年 1 月开始，任何已发表成果的科研人员和机构均可使用该服务。该服务专注于元数据的管理（代表科研人员起草完整、准确和适当的元数据，使其遵从 FAIR 原则），同时防止敏感数据以不适当的格式发布以及使用不适当的许可协议。敏感数据的发布和共享其实是一个挑战，尽管服务人员会根据数据匿名化的既定准则检查作者的提交内容，并在适当的情况下要求作者提供参与者同意的详细信息以发布数据，但是，承担责任的仍旧是作者本人。因此，最好的方法是建议作者将“非匿名化数据”存储到托管或限制访问的受保护存储库中，如 UK Data Archive。

希望使用服务的科研人员需填写服务查询表格，见表 5，提供数据特征，检查数据是否符合服务标准（服务标准如《可以帮助组织和分享的研究数据类型》[28]、《科研数据支持的提交、检查和输出》[29]、《科研数据支持条款和条件》[30]）。服务期间，数据管理过程中所有耗时的工作均有专业的数据编辑团队负责，具体包括：①查找数据存储库并提供遵守政策的建议；②提供数据可用性声明编写建议以及数据引用建议；③上传数据至 figshare，最高可达 50GB；④检查是否存在敏感数据；⑤确保最终用户可以获取数据；⑥确保数据文件与元数据和相关出版物相匹配；⑦将文件以结构化方式组织和收集；⑧优化提交的元数据（包括标题和数据集的描述）；⑨根据行业标准定义对内容进行分类；⑩以列表形式展现作者和相应论文；⑪为每个数据集生成 DOI；⑫数据集匿名访问设置；⑬链接并同步数据到相关论文；⑭提供评价报告以及数据编辑的反馈意见；⑮当数据存储量超过 50GB 时，按要求支付费用获取额外存储空间。付费服务的优势体现在三个层面（见表 6），享受服务的科研人员需支付 300 欧元/ 265 英镑/ 340 美元的服务费用（不含增值税和地方税）。该服务目前已经帮助出版了包括《自然》杂志在内的 Springer Nature 期刊，BMC 生态学等多种期刊的百余个数据集。

表 5　付费服务查询表格[27]

数据信息	姓名
	邮件
	数据是否与手稿（研究论文、文章、章节或专著）有关
	相关稿件是否发表
服务选择	同意与数据支持服务进行联系
	同意与数据支持服务进行联系，同时接收 Springer Nature 营销通讯

表 6　Springer 付费数据服务的优势

科研人员	• 由经验丰富的数据编辑人员进行专业管理，确保数据容易地被找到和重用 • 节省存储、共享和管理数据的时间，以科学方式组织和呈现数据 • 增加团队在研究领域的曝光率 • 更轻松地满足出版商、所属机构和资助机构的数据共享要求 • 使共享数据的科研人员成为开放科学的先驱
科研工作	• 通过发布科研数据，并将其链接到相应论文，提高研究内容的引用率 • 确保与论文相关数据可用且可查找，提高期刊文章的可见度
研究领域	• 通过出版科研数据，推动变革，激励并推动科学创新 • 使数据更容易被获取和发现，提高研究的透明度和可重复性

实际上，该服务已于 2018 年初结束了试验阶段，并收取了一定费用以确保服务的可持续性，但服务目前仍在继续开发和测试，比如正在考虑为会议提供数据管理支持。需要说明的是，像 Springer Nature 这样，由学术出版商提供科研数据管理支持是科研数据管理服务的一项新发展。以往，科研人员和期刊编辑往往是通过其机构内的科研数据管理团队，通过咨询科研资助机构，通过专业的数据监管机构，如 Jisc 或英国的数字监管中心（Digital Curation Centre，DCC），或者直接与专业数据存储库合作来实现上述目标。但是，并非所有的科研人员都可以通过这些途径获得服务，因此，

未来的科研数据管理服务，一定是一种混合的模式，即由不同利益相关群体提供的服务进行组合的方式，没有统一的最佳方案。由出版商提供科研数据管理服务是机构科研数据管理服务的有益补充，特别是当机构由于人员、资源限制无法不断扩展或提供足够灵活的服务时。

4　Springer Nature 政策标准化框架应用情况及其启示

4.1　Springer Nature 政策标准化框架应用情况

该政策框架是在 2016 年第二季度引入 Springer Nature 期刊，其中包括《自然》，BioMed Central 期刊以及 SpringerLink 上托管的期刊[31]。截至 2019 年 9 月，Springer Nature 出版集团旗下已有 1 600 多种期刊基于框架提出了自己的数据政策。经过人工识别，将已提出政策的 1 600 多种期刊进行学科分类，发现自然科学类期刊占比最高，为 71.5%，其中生命科学类期刊采用类型二和类型三政策的居多。其次是交叉科学类期刊，占 15.5%，社会科学类期刊最少，仅有 13%。以社会科学领域为例，截至 2019 年 9 月，类型一采用最多(占 72.4%)，其次是类型二(占 21.7%)，类型三(占 4.5%)，另外还有 1.4% 的社科类期刊还处于未确定类型，状态为政策类型待选。研究发现，除了学科因素外，其他一些因素也会影响期刊政策类型的选择。比如，如果编辑部人员不足，难以执行政策执行情况检查的期刊可能会选择类型一或类型二；之前没有数据共享政策的期刊也更可能采用类型一或类型二。类型四要求每个出版物都必须开放数据，这是目前采用最少的政策类型，迄今为止仅被由高度重视数据共享的数据期刊和/或具有完善开放科学文化和开放数据共享知识库的作者社区采用。随着科研数据共享政策强度的增加，发布过程中的成本也会随之增加。类型三和类型四要求对已提交或已接受的手稿进行额外的检查，这可能会产生相关成本，对于期刊和出版商来说，了解这些成本非常重要。

4.2　Springer Nature 推进政策标准化工作的启示

4.2.1　相关政府部门应牵头出台科研数据政策框架

在科学数据的开放共享体系中，出版商是关键利益相关者。出版商与资助机构、数据中心与存储库以及科研机构等其他利益相关者合作，共同促进科研数据开放共享文化的传播。尽管目前我国的出版机构已经意识到科研数据开放共享的重要价值，但与国外出版商相比，我国的科技期刊出版机构科研数据共享起步晚，无论是期刊还是作者，对于共享科研数据的益处还没有充分认识到，对科研数据共享政策重视程度都不够。出版机构参与开放共享科研数据的实践还未形成规模，快速制定有效的、具有可操作性的科研数据政策，面向科研数据提交、审查、存储和引用等关键环节给出明确而具体的说明、建议或者规定，并明确相应的监督检查方式，还有较大难度。因此相关政府部门，应充分发挥我国的制度优势，尽快建立起我国科研数据共享的法律体系，明确规定科研数据共享的知识产权以及恶意使用数据必须承担的法律后果。与此同时，建议作为期刊管理部门的中国科协牵头召集科研数据共享的利益相关群体，讨论制定并推出面向科技期刊的科研数据政策框架，以指导和帮助期刊尽快顺利推出适合期刊自己个性化需求的、有效的且符合国际最佳实践的科研数据政策。

4.2.2　借鉴国际先进经验建构科学合理的政策框架

国家创新能力根植于知识创造、汇聚与传播及其生态环境。科技期刊传承人类文明，荟萃科学发现，引领科技发展，一定程度上它是国际科技竞争力和文化软实力的体现[32]。2018 年 11 月 14 日召开的中央全面深化改革委员会第五次会议审议通过了《关于深化改革培育世界一流科技期刊的意见》。采用国际通用语言是建设世界一流期刊不可回避的重要路径。随着科研数据共享和可获取呼声的高涨，国际上越来越多的期刊开始重视科研数据的存缴和出版，纷纷制定期刊自己的科研数据政策，要求作者在发表论文时公开相关原始数据[33]。作为期刊大国，打造有影响力的世界一流科技期刊，必须顺应学术交流发展态势，强化科技期刊建设的顶层设计，发挥政策引导作用，推动开放科学生态环境构建。与国外相比，国内拥有数据政策的期刊数量少，期刊数据政策研究没有引起国内相关机构和学者足够的重视。为此，我国的科技期刊管理机构应该积极借鉴国际经验和实践，在界定科研数据共享的利益相关群体角色、责任和义务时，可以考虑借鉴 Elsevier 的数据生产者(科研人员)要求、利益相关者责任以及数据引用者义务三个维度，细化数据所有者权益并明确义务；可以借鉴 Springer Nature 确定的科研数据政策九个共有要素，广泛咨询并听取我国科技期刊、图书馆员和科研资助机构的意见，结合学科的差异性，建构包含若干类型的、通用的科技期刊政策框架，指导帮助国内期刊改善已有数据政策，促进期刊数据政策全面系统的发展，提升期刊数据政策的利用效率，使之更加科学合理。

4.2.3　设计配套支持服务有效推动政策框架的应用

在“建设世界一流科技期刊”的背景之下，科技期

刊和社科期刊都应该关注最新的出版技术、出版工具和出版模式,并积极尝试,努力顺应国际期刊的发展潮流,提升期刊的工作效率和传播效果[34]。无论对于期刊编辑部还是科研人员而言,科研数据出版都属于新兴事物,不管是认知水平和知识储备上,还是在实践操作上都与期望的最佳实践存在较大差距,因此必须设计配套的支持服务以有效推动政策框架的应用。Springer Nature 同时出台的面向科研人员的配套支持服务和面向期刊编辑部的配套支持服务给我们提供了很好的示范。但是,有一点必须注意的是,我国无论是公共的综合型数据知识库还是学科领域的数据知识库与国外相比在数量上都存在较大差距,如果期刊推出科研数据政策,就必然会面临巨大的维护数据的成本和负担,这会导致很多期刊望而却步。因此,必须由政府机构牵头来推动科研数据基础设施的建设,但是需要注意的是科研数据基础设施建设是一项系统工程,其服务能力不取决于单一项目或系统的规模、性能和指标,而需要一套科学的长效运行机制[35],建设的目标不是构建一个超级数据存储中心,而是构建一个可以分享数据、促进研究者合作与交流,从而加速创新,为发展提供动力的数据基础设施[36-37]。只有依托科研数据基础设施开展配套服务,才能真正推动服务的发展,体现出服务的价值和意义,推动服务不断深入和完善,进而顺利实现框架的应用,促进政策的制定、完善和有效实施,达到开放科研数据的预期成效。

参考文献:

[1] SCHMIDT B, GEMEHOLZER B, TRELOAR A. Open data in global environmental research: the Belmont Forum's open data survey[J]. PLOS ONE, 2016,11(1):e0146695.
[2] CENTER for OPEN SCIENCE. Transparency and openness promotion guidelines[EB/OL]. [2020-03-10]. https://cos.io/top/.
[3] JONES L, GRANT R, HRYNASZKIEWICZ I. Implementing publisher policies that inform, support and encourage authors to share data: two case studies[EB/OL]. [2020-03-10]. https://insights.uksg.org/articles/10.1629/uksg.463/.
[4] NAUGHTON L, KERNOHAN D. Making sense of journal research data policies[J]. Insights the UKSG journal, 2016,29(1):84-89.
[5] TAICHMAN DB, SAHNI P, PINBORG A, et al. Data sharing statements for clinical trials[EB/OL]. [2020-03-10]. https://www.bmj.com/content/bmj/357/bmj.j2372.full.pdf.
[6] COLAVIZZA G, HRYNASZKIEWICZ I, STADEN I, et al. The citation advantage of linking publications to research data[EB/OL]. [2020-03-10]. https://arxiv.org/abs/1907.02565.
[7] UK Research and Innovation. Common principles on data policy [EB/OL]. [2020-03-10]. https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/.
[8] COLAVIZZA G, HRYNASZKIEWICZ I, STADEN I, et al. The citation advantage of linking publications to research data[EB/OL]. [2020-03-10]. https://arxiv.xilesou.top/pdf/1907.02565.pdf).4.
[9] GIOFRE D, CUMMING G, FRESC L, et al. The influence of journal submission guidelines on authors' reporting of statistics and use of open research practices[J]. PLOS ONE, 2017(12): e0175583.
[10] VASILEVSKY NA, MINNIER J, HAENDEL MA, et al. Reproducible and reusable research: are journal data sharing policies meeting the mark? [J]. Peer J, 2017(5):e3208.
[11] HRYNASZKIEWICZ I, BIRUKOU A, ASTELL M, et al. Standardising and harmonising research data policy in scholarly publishing[EB/OL]. [2020-03-10]. http://ijdc.net/index.php/ijdc/article/download/12.1.65/467.
[12] HRYNASZKIEWICZ I, SIMONS N, HUSSAIN A, et al. Developing a research data policy framework for all journals and publishers[EB/OL]. [2020-03-10]. https://datascience.codata.org/article/10.5334/dsj-2020-005/.
[13] 王丹丹. 英国《开放科研数据协议》的解读与启示[J]. 数字图书馆论坛,2018(8):25-30.
[14] SPRINGER NATURE. Research data policy types[EB/OL]. [2019-12-17]. https://www.springernature.com/de/authors/research-data-policy/research-data-policy-types.
[15] SPRINGER NATURE. Helpdesk[EB/OL]. [2019-12-17]. http://www.springernature.com/gp/group/data-policy/helpdesk.
[16] ASTELL M, HRYNASZKIEWICZ I, GRANT R, et al. Have questions about research data? Ask the Springer Nature Helpdesk[EB/OL]. [2020-03-10]. https://figshare.com/articles/Providing_advice_and_guidance_on_research_data_a_look_at_the_Springer_Nature_Helpdesk/5890432/2.
[17] SPRINGER NATURE. Recommended repositories[EB/OL]. [2020-03-10]. http://www.springernature.com/gp/group/data-policy/repositories.
[18] SPRINGER NATURE. Data availability statements[EB/OL]. [2020-03-10]. http://www.springernature.com/gp/group/data-policy/data-availability-statements.
[19] NATURE. Announcement - where are the data? [EB/OL]. [2020-03-10]. http://www.nature.com/news/announcement-where-are-the-data-1.20541.
[20] GRANT R, HRYNASZKIEWICZ I. The impact on authors and editors of introducing data availability statements at Nature journals [EB/OL]. [2020-03-10]. http://www.ijdc.net/article/view/614.
[21] SPRINGER NATURE. Research data support service[EB/OL]. [2019-12-17]. https://www.springernature.com/de/authors/research-data/research-data-support.
[22] SPRINGER NATURE. Research data at springer nature[EB/OL]. [2019-12-17]. https://researchdata.springernature.com/?

utm_source = springernature. com&utm_campaign = CMTL_USG_SOOD19.
[23] STUART D, BAYNES G, HRYNASZKIEWICZ I, et al. Practical challenges for researchers in data sharing[EB/OL]. [2020－03－10]. https://figshare. com/articles/Whitepaper_Practical_challenges_for_researchers_in_data_sharing/5975011/1.
[24] MARKET RESEARCH. Wiley open science researcher survey 2016 [EB/OL]. [2020－03－10]. https://figshare. com/articles/Wiley_Open_Science_Researcher_Survey_2016/4748332/2.
[25] BERGHMANS S, COUSIGN H, DEAKIN G, et al. Open data: the researcher perspective - survey and case studies[EB/OL]. [2020－03－10]. https://data. mendeley. com/datasets/bwrnfb4bvh/1.
[26] HRYNASZKIEWICZ I. New services to support open research [EB/OL]. [2020－03－10]. https://blogs. biomedcentral. com/bmcblog/2017/04/27/new-services-to-support-open-research/.
[27] SPRINGER NATURE. Research Data Support service[EB/OL]. [2020－05－31]. https://www. springernature. com/gp/authors/research-data/research-data-support.
[28] SPRINGER NATURE. Types of research data we can help to organise and share[EB/OL]. [2019－12－17]. https://media. springernature. com/full/springer-cms/rest/v1/content/15403688/data/v1.
[29] SPRINGER NATURE. Organise and share your data faster with research data support[EB/OL]. [2019－12－17]. https://media. springernature. com/full/springer-cms/rest/v1/content/15403690/data/v3.
[30] SPRINGER NATURE. Please read these terms and conditions carefully before using the site and ordering the services[EB/OL]. [2019－12－17]. https://media. springernature. com/full/springer-cms/rest/v1/content/15541262/data/v1.
[31] HRYNASZKIEWICZ I. Promoting research data sharing at springer nature[EB/OL]. [2020－03－10]. http://blogs. nature. com/ofschemesandmemes/2016/07/05/promoting-research-data-sharing-at-springer-nature.
[32] 中国科协 中宣部 教育部 科技部关于深化改革培育世界一流科技期刊的意见[J]. 分析化学,2019,47(9):1301,1365,1381,1426.
[33] 傅天珍,陈妙贞. 我国学术期刊数据出版政策分析及建议[J]. 中国出版,2014(23):31－34.
[34] 武晓耕,胡小洋,占莉娟,等. “建设世界一流科技期刊”背景下的社科期刊发展策略探究[J]. 中国科技期刊研究,2019,30(8):801－805.
[35] 章昌平,米加宁,黄欣卓. 超越数据洪流:第四研究范式下的社会科学研究数据基础设施[J]. 学海,2019(3):55－61.
[36] 王敬,王彦兵. 国外科研数据基础设施研究及实践的调研与分析[J]. 情报资料工作,2016(6):99－104.
[37] 马骏. 科学数据管理服务推进过程中的合作策略[J]. 情报资料工作,2018(05):62－67.

作者贡献说明:
王丹丹:提出研究思路,论文撰写及最终版本修订;
刘清华:论文撰写、修改;
葛力云:协助进行论文数据收集和整理。

Springer Nature's Practice for Research Data Policy Standardization and Its Inspiration

Wang Dandan　Liu Qinghua　Ge Liyun
College of Management, Henan University of Science and Technology, Luoyang 471023

Abstract: [**Purpose/significance**] Springer Nature has made pioneering contributions in promoting the standardization and unified development of research data policies. Study the standard framework of research data policy proposed by Springer Nature and its application, hoping provide a reference for China's sci-tech periodicals management agencies. [**Method/process**] Through web search, text analysis, website visits, and comparative analysis, this paper analyzed Springer Nature's policy standardization framework, researched on the support services provided by Springer Nature to promote the application of the framework. Then it summarized Springer Nature's work practices of promoting the standardization systematically and revealed its significance for reference. [**Result/conclusion**] Springer Nature's policy standardization framework can be used as a reference for China's relevant departments to formulate a standard framework for research data policies. Third-party data curation service developed by Springer Nature to promote the standardization framework, shows the practical path for publishers to participate in research data management services. It is suggested that China's sci-tech periodical management organizations take the lead to convene stakeholder groups, draw on international advanced experience to construct a scientific and reasonable policy framework, and design supporting services to effectively promote the application of the policy framework.

Keywords: publisher　research data　policy　service